Negative and Positive *cis*-Acting Elements Control the Expression of Murine α_1 -Protease Inhibitor Genes

K. T. MONTGOMERY,¹ J. TARDIFF,² L. M. REID,¹ AND K. S. KRAUTER^{2*}

*Departments of Cell Biology² and Molecular Pharmacology,¹ Albert Einstein College of Medicine,
1300 Morris Park Avenue, Bronx, New York 10461*

Received 29 August 1989/Accepted 14 February 1990

The α_1 -protease inhibitor (α_1 -PI) proteins of mice are encoded by a group of genes whose members are expressed coordinately in a liver-abundant pattern and are regulated primarily at the transcriptional level. To better understand the developmental and tissue-specific regulation of this gene family, one member that is analogous to the human α_1 -antitrypsin gene was chosen for study. Deletional analysis of the upstream regulatory region of this gene was performed, spanning from -10 kilobases to -80 base pairs relative to the transcriptional start site. Two functional positive *cis*-acting elements within the 522 bases immediately upstream of the start site for transcription were shown to modulate the level of expression from this promoter when introduced into human or mouse hepatoma cells, and a third region acted as a negative regulatory element in that its deletion resulted in a two- to sixfold increase of expression of a transfected minigene construct. Sequence comparison between the regulatory domains of two mouse α_1 -PI genes and the human α_1 -antitrypsin gene showed that the mouse gene contains a novel positive *cis*-acting element which is absent in human gene and that a specific eight-base-pair difference between species results in a strong positive *cis*-acting element in the human gene acting as a negative element in the mouse gene. An enhancer located ~3,000 base pairs upstream of the major start site for transcription was also identified. This element is position and orientation independent. Several different DNA-protein binding assays were used to demonstrate that each DNA segment with functional significance in transfection assays interacts specifically with proteins found in adult mouse liver nuclei. The major positive-acting element appeared to be specifically recognized by nuclear proteins found only in tissues that express α_1 -PI, while the negative element binding proteins were ubiquitous. Thus, the distal regulatory domain including bases -3500 to -133 of this murine α_1 -PI gene family member is more complex than was previously demonstrated. It is composed of a set of at least three additional functional *cis*-acting regulatory elements besides those which have been mapped by others and has a far upstream enhancer.

In the mouse, there are five closely related α_1 -protease inhibitor (α_1 -PI) genes which are located on chromosome 12 (22, 29; F. Borriello and K. S. Krauter, submitted for publication). These genes are homologous to the human serine protease inhibitor α_1 -antitrypsin (α_1 -AT). All five genes are coexpressed in a liver-abundant pattern and have been shown to be regulated at the transcriptional level during development and differentiation (7, 41). The primary function of human α_1 -AT is to inactivate neutrophil elastase in the lung. Genetic defects which result in reduced serum levels of α_1 -AT can result in severe lung and liver disease (13). Therefore, we and others have sought to understand the regulation of the murine gene family as a prelude to the establishment of a mouse model system in which to study the pathology associated with α_1 -AT deficiency. Two of the five α_1 -PI genes which are expressed in the adult mouse liver (Borriello and Krauter, submitted) encode the critical Met-Ser dipeptide in exon 5, thought to be important for normal antielastase activity in the human lung (23, 45). However, the specific target proteases and the physiological roles of the multiple murine protease inhibitor gene products are not yet established.

It has been reported that low-level liver-specific transcription of one murine α_1 -PI gene requires just 168 base pairs (bp) upstream of the cap site, while high-level transcription depends on an enhancer located between positions -522 and -168 relative to the transcriptional start site (18). Competi-

tive in vitro binding studies and crude deletion mapping have suggested that several regulatory proteins known to bind to the regulatory regions of other liver-specific genes interact with DNA sequences in this region. More specifically, the liver-abundant transcription factor CAAT/enhancer-binding protein (C/EBP) (24) is able to bind at positions -450 and -200 upstream of the start site, and another hepatic nuclear DNA-binding protein, hepatic nuclear factor 3 (HNF 3), binds at -190 and -370 (3, 17, 18).

To delineate further the elements required for transcription of the mouse α_1 -PI gene family, we performed a detailed deletional analysis of one cloned gene. Here we report that the α_1 -PI-I gene contains functional *cis*-regulatory elements which have not previously been reported. Three of these elements are located between -318 and -133 relative to the transcriptional start site, and one lies ~3 kilobases (kb) upstream. Two of the promoter-proximal elements act to enhance transcription from the α_1 -PI-I promoter, while the third inhibits it when transfected into human and mouse hepatoma cell lines. We demonstrated that mouse liver nuclear proteins interact specifically with DNA sequences within each of the four functionally defined regions. Furthermore, our analysis showed that the deletion of some of the DNA sequences in this upstream region which have been shown by others to bind to specific proteins has little effect on expression in a transient assay system.

Finally, we showed that one other mouse α_1 -PI gene which is expressed in adult liver contains all the same critical DNA sequences within the promoter region as the gene we

* Corresponding author.

analyzed. Such high conservation of DNA sequences in a flanking region argues either that these sequences are important in gene regulation or that these two genes are the result of a very recent duplication. In contrast, a comparison of the sequence of the two murine genes from -522 to +44 with the published sequence of the human α_1 -AT gene (34) reveals portions which are highly conserved and other segments which have diverged. We show evidence that the highly conserved regions endow both murine and human genes with basal levels of expression in hepatic cells, while divergence in the region identified as a strong enhancer of transcription of the human gene (8) may contribute to species differences in expression.

MATERIALS AND METHODS

Cell culture and transfection. Cells were maintained as monolayers in a 1:1 mixture of Dulbecco modified Eagle medium and Ham F12 medium (GIBCO Laboratories, Grand Island, N.Y.) plus 10% fetal calf serum, penicillin (50 U/ml), streptomycin (50 μ g/ml), and dexamethasone (10^{-7} M). They were passaged 18 h before transfection and plated at the density which would result in a 50% confluent plate at the time of transfection. The cell lines used were human hepatoma cell line HepG2 (26), mouse hepatoma cell line BW73 (48), human neuroblastoma cell line SK-N-SH/SY5Y (2), and murine fibroblast line NIH 3T3.

Covalently closed circular plasmid DNA was transfected into cells by CaPO_4 -mediated DNA transfer by standard procedures (16). Preliminary experiments to determine the capacity of each cell type for uptake and expression of DNA were performed, and the total amount of DNA per cell was adjusted to ensure linearity of signal to input DNA. For the experiments shown, transfection included 3 μ g of the control construct, mouse metallothionein-I-human growth hormone-releasing factor (MTI-hGRF), and 30 μ g of the -522-hGRF plasmid or equivalent molar amounts of other constructs. The control construct was expressed at high levels in all cells tested.

Clones and constructs. All the α_1 -PI DNAs used were obtained from BALB/c genomic clones gliv3N (α_1 -PI-I) and gliv3B (α_1 -PI-II) described previously (29). The control plasmid, MTI-hGRF, was provided to us by R. Evans of the Salk Institute and contains the murine MTI promoter along with the first 64 bases of exon 1 linked to exon 1 of a hGRF minigene reporter sequence at an engineered *Bgl*II site (19). The -522-hGRF construct was made by replacing the MT promoter fragment in the MTI-hGRF plasmid with 537 bases of α_1 -PI-I sequence, beginning at an *Eco*RI site at position -522 of the α_1 -PI-I gene and ending at a *Bst*EII site located at +15. This construct produces an mRNA which encodes a functional hGRF protein and differs from the mRNA produced by the control construct only in the 5' untranslated region.

5'-end deletion mutants were derived from the -522-hGRF construct by linearizing the circular plasmid by digestion with *Eco*RI and then treating the DNA for increasing times with exonuclease III (21). Fragments containing the deleted promoter DNA and the entire hGRF minigene were excised by digestion with *Hind*III and then reinserted into the cloning vector pGEM-1 (Promega Biotec, Madison, Wis.). All constructs contained identical plasmid vector sequences adjacent to test inserts. The endpoint of each deletion mutant was determined by DNA sequence analysis (28).

Internal deletion plasmids were made from the 5'-end

deletion constructs by ligating upstream fragments to selected deletion constructs. The upstream fragments were generated with selected synthetic oligonucleotide primers in the polymerase chain reaction (PCR) (10) with a template which included the entire α_1 -PI upstream region. The upstream PCR primer contained a *Bgl*II restriction enzyme site linked to the 10 nucleotides immediately downstream of, but not including, the *Eco*RI site at position -522 in the -522-hGRF construct. The 3'-end primer was chosen to include the 10 nucleotides immediately upstream of the 5' endpoint of the desired internal deletion plus an additional 6 nucleotides corresponding to an *Eco*RI restriction endonuclease site. The PCR products were directionally cloned into the *Bgl*II-*Eco*RI site in the polylinker of pSP72 (Promega). Then the appropriate 5'-end deletion mutant insert which included a truncated α_1 -PI-I promoter sequence linked to the hGRF reporter was directionally cloned into the *Eco*RI-*Hind*III site immediately 3' of the PCR insert. This procedure resulted in (i) the deletion of the selected nucleotides between the 3' endpoint of the PCR-generated upstream fragment and the 5' end of the selected 5'-end deletion construct and (ii) the insertion of the 16-mer GAATTCGAGCTCGCCC at the point of the deletion owing to remaining polylinker DNA. A construct which included 10 kb of upstream sequence was generated by inserting the 9.5-kb *Eco*RI fragment from genomic clone gliv3N at the native *Eco*RI site retained at -522 in the α_1 -PI-hGRF construct, in the orientation which restored the normal genomic flanking sequence. Smaller constructs used to localize the far upstream enhancer region were made by subcloning appropriate restriction fragments derived from the upstream region of gliv3N into the unique *Eco*RI site located at the 5' end of deletion constructs -340-hGRF and -133-hGRF. The nucleotide positions of the fragments are indicated in the figures.

RNA isolation and quantitation of mRNA. RNase protection was performed essentially as described by Melton et al. (36). Cytoplasmic RNA was isolated from cultured cells by Nonidet P-40 lysis and phenol-chloroform-isoamyl alcohol (50:49:1) extraction (44). Cytoplasmic RNA (30 μ g) isolated from transfected cells was mixed with 1×10^6 cpm of single-stranded riboprobe labeled to greater than 5×10^8 cpm/ μ g with [32 P]GTP and ethanol precipitated. The precipitate was recovered and suspended in 10 μ l of hybridization buffer (80% formamide, 40 mM PIPES [piperazine-*N,N'*-bis(2-ethanesulfonic acid)] pH 6.7, 400 mM NaCl, 1 mM EDTA), covered with paraffin oil, heated to 85°C for 10 min, and then incubated at 72°C for 3 to 4 h. After hybridization, samples were digested with RNase A (100 μ g/ml) and RNase T₁ (20 U/ml), the digestion was stopped by the addition of sodium dodecyl sulfate and proteinase K and incubation at 37°C for 15 min, and samples were phenol-chloroform-isoamyl alcohol extracted, ethanol precipitated, and subjected to polyacrylamide gel electrophoresis with 8% polyacrylamide-8 M urea in a standard DNA-sequencing gel apparatus. Gels were dried and subjected to autoradiography without intensifying screens followed by scanning with an LKB Ultrascan. Exposures were adjusted to ensure linear film response.

Preparation of crude nuclear extracts. Crude nuclei were prepared from mouse liver essentially as described previously, omitting the high-sucrose final purification step (7). Nuclear pellets were then subjected to high-salt extraction with 0.36 M NH_4SO_4 exactly as described by Parker and Topol (39). Protein extracts contained greater than 3 mg/ml.

Gel retardation and DNA footprint analyses. Three different methods were utilized to determine the positions of

protein interactions with specific DNA sequences. In all cases, 3'- or 5'-end-labeled DNA restriction fragments were prepared as described previously (29). For both methylation interference studies and copper footprinting analyses, DNA-protein complexes were isolated by using a low-salt gel retardation step (12, 15). Briefly, $\sim 1.5 \times 10^5$ cpm (10 to 20 ng) of end-labeled probe either methylated by dimethyl sulfate treatment (30 s in 1:200 dilution) (35) or untreated was mixed with 100 μ g of poly(dI-dC) per ml in 20 mM Tris hydrochloride (pH 8.0), 70 mM KCl, 5 mM MgCl₂, 0.5 mM CaCl₂, 0.5 mM dithiothreitol, 0.1 mM EDTA, and 7% glycerol. Crude nuclear extract (10 to 15 μ g) was added per nanogram of DNA in a total reaction volume of less than 100 μ l, and the mixture was incubated for 20 min at 25°C. Samples were subjected to gel retardation analysis and then exposed to X-ray film to locate the position of bands.

For methylation interference studies, bands were excised from the gel, the gel slices were crushed, and the DNA was eluted in 0.5 M NH₄COOH and 0.1 mM EDTA and recovered by ethanol precipitation. DNA was suspended in freshly prepared 10% piperidine, incubated at 90°C for 40 min to cleave at methylated G residues, lyophilized to dryness, washed with H₂O and dried, and then suspended in 90% formamide, heated to 90°C for 5 min, and subjected to gel electrophoresis on 6% DNA-sequencing gels.

Organic copper footprinting (30) was performed essentially as described by Morrow et al. (38). DNA fragments were subjected to gel retardation analysis as described above. After electrophoresis, the gel was immersed in 50 mM Tris hydrochloride (pH 8.0), 1.66 μ M 1,10-phenanthroline (Sigma Chemical Co., St. Louis, Mo.), 0.38 mM CuSO₄, and 4.8 mM 3-mercaptopropionic acid for 20 min at 25°C with gentle agitation. The reaction was quenched by the addition of 2.15 mM 2,9-dimethylphenanthroline. Labeled fragments were then identified by autoradiography, eluted from the gel, and subjected to gel electrophoresis as described above.

DNase I footprinting with crude nuclear extracts was performed essentially as described by Lichtsteiner et al. (33). Briefly, 1 ng of end-labeled DNA fragment was added to a solution which had been preincubated for 10 min on ice containing 25 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid) (pH 7.6), 60 mM KCl, 5 mM MgCl₂, 1 mM EDTA, 7.5 mM dithiothreitol, 5 mM phenylmethylsulfonyl fluoride, 500 μ g of poly(dI-dC) per ml, 8% glycerol, and 2 mg of crude nuclear lysate per ml in a final volume of 100 μ l. After 30 min on ice, 100 μ g of DNase I per ml was added and incubation continued for 15 min. The reaction was terminated by the addition of 4 volumes of 20 mM Tris hydrochloride (pH 7.5), 20 mM EDTA, 0.5% sodium dodecyl sulfate, 75 μ g of salmon sperm DNA per ml, and 50 μ g of proteinase K per ml. After incubation for 60 min at 45°C, the DNA was extracted with phenol-chloroform-isoamyl alcohol, ethanol precipitated, and electrophoresed as described above.

RESULTS

Sequence comparison between mouse and human α_1 -PI genes. We have begun analysis of the regulatory domains of two murine α_1 -PI genes, designated here as α_1 -PI-I and α_1 -PI-II, which are very similar to the single active human α_1 -AT gene. Both mouse genes encode the critical Met-Ser dipeptide in their exon 5 active sites (23, 29) and are expressed at similar high levels in mouse liver (Borriello and Krauter, submitted). The transcriptional start site for the

murine α_1 -PI-I liver transcript lies within a 4.3-kb *Eco*RI fragment (17). Exon 1 is 44 bases long and is located ~ 6 kb upstream of exon 2.

In Fig. 1, we show partial sequences of the two mouse genes aligned with the human α_1 -AT gene sequence (34) in the region of the transcriptional start site. Two general features emerge from this comparison. (i) Greater than 98% sequence identity exists between the two mouse genes despite the fact that the sequences lie largely in an untranscribed 5'-flanking region. (ii) Variable levels of sequence conservation exist when one compares the human gene with the mouse genes through the 5' region and in exon 1. The identity between murine and human sequences in exon 1 and the proximal upstream region between -122 and +44 is >75%, while the overall identity drops to <58% between -522 and -122.

Two liver-specific nuclear factors originally isolated from rat liver nuclear extracts have been shown to bind to the human α_1 -AT gene in the proximal upstream region and endow it with a basal level of expression in hepatic cells. These factors have been designated LF-A1 (8, 37) or HNF II (32) and LF-B1 (8, 37) or HNF I (6). In the human gene, they bind at nucleotides -125 to -101 and -82 to -52, respectively. Sequence comparison showed that the HNF I-binding site is highly conserved between the mouse and human genes, with respect to both DNA sequence and position relative to the start site of transcription (Fig. 1). The HNF II site is less well conserved but is discernable. Expression data presented below suggest that the proximal 80 bp of the mouse gene analyzed here are sufficient to provide for tissue-specific expression, while it has been shown that the proximal 137 bp of the human gene (8, 37) provide for tissue specificity.

Our functional analysis focused on the distal regulatory domain between -133 and -522 where the mouse-human divergence is greatest. Experimental evidence showed that despite the fact that the mouse gene studied encodes a protein very similar to the human α_1 -AT protein, their functional regulatory elements may differ. This divergence between the species in the distal regulatory region may contribute to the observed differences in the tissue distribution and control of expression of the human and mouse genes which recent transfection (8, 18, 42) and transgenic (25, 27, 43) studies have revealed.

5' deletion analysis of α_1 -PI-I upstream region. It has been shown by transfection analysis that the region from -522 to -168 contains *cis*-acting sequences which behave as a functional enhancer of the α_1 -PI-I gene (18). We extended this observation by a more detailed deletion analysis of this region which defined additional functional elements. The strategy used is outlined in Fig. 2. Briefly, a minigene was prepared in which 537 bp of α_1 -PI sequence (-522 to +15) were linked to a reporter DNA sequence encoding hGRF. This reporter was chosen because its transcription was known to produce high levels of stable hGRF mRNA in hepatocytes and its expression could be readily detected by nuclease protection assay (19). A set of deletion constructs with 5' endpoints between -522 and -80 was prepared and transfected into a variety of cell lines. As an internal control for transfection efficiency, the mouse MTI gene promoter linked to the same reporter sequence was cotransfected with test constructs. Transcripts driven by the α_1 -PI promoter protected a 120-nucleotide fragment of a ³²P-labeled antisense RNA probe in an RNase protection assay, while MTI promoter-driven transcripts protected a 105-nucleotide fragment of the same probe.

DNA Sequence Comparison of Mouse and Human Protease Inhibitor Regulatory Regions

Mouse PI-II	-522	-507	-492	-477	-462	-447	-432	-417	120
Mouse PI-I	GAATTCATATCTACA	CTTCCAAAGGAGCAAT	CTCTTGTGTTTGGTC	CTGTGCTCCATTCCG	TCAGTGGGCGACATA	ACCTACTCTGAGCCC	CTGGTGAGGCGAGTG	AATCTTCT*CCCCAG	119
Human AAT	AGGC G CTGGGA	TGG T	GG	CC C CC AG	CT	AAC T *	T T	CA * C A G	118
Mouse PI-II	-403	-388	-373	-359	-346	-332	-317	-302	236
Mouse PI-I	GCATCTCGAAGCAGA	TGGCTCAGTAAATAT	T*GACTTTGCTTACA	TGCAAGCAACCC**A	AGCATTCCCCAG*GA	GAGTGCAGGACCAAGC	CAACCATGTTTGTC	TTCCTAGAGGCCATG	235
Human AAT	T T GG **** A	GG A TC	AC	GTGG A G ACT	G	TG T	CAG G	T **** G ACT C	230
Mouse PI-II	-287	-272	-257	-242	-227	-212	-197	-182	356
Mouse PI-I	TGACTCATTACACCA	GAACGCTTTGGACA	CCGATGTTTGTTT	GTGTGTTTATGAGG	GATGGTCTTCTAAC	TGCTTGGCTTAAGAC	TCCATTGATTAGGA	GCAGAGAGAGCTGAG	355
Human AAT	<-AP1->	CGC C CTC A C	A C C	*** *****	C C **	ACAA G C	CGG ** G GA**	CT T C CT CC A GC	336
Mouse PI-II	-167	-152	-138	-123	-108	-93	-78	-63	474
Mouse PI-I	CTCTGTCCCGAACAG	TGTGGGTGGTAC*CT	CAGTGCOCACATCGAT	GAACCTTAGCCCTCT	CTGTTTGGCTTACCA	ATAACTGTGGTGACA	TGTTTAATATTCAC	TAGCAAACTCCOCCA	474
Human AAT	AAAGGGT GG	C A C GCGA	AT GC AG G	*** *	CT G	CT G	C	C GC	452
Mouse PI-II	-48	-33	-18	+1	+13	+27	+44		591
Mouse PI-I	TATCCOCTTGGCTC	CACTGCTTAATAC	AGACTAGGACAGGCG	TCTGTCTCTCAGCC	TCGGTCACCACC*CA	GCTCTGGGACAACAA	GCTGTGAGTATATT	TTCCTGGGAAATA	593
Human AAT	TG TC A	*	G G	C	T A G	A T AC	GTG ATC A	GCC A C AGGGGT	571

FIG. 1. Alignment of the nucleotide sequence of the upstream regions of two murine α_1 -PI genes with that of the human α_1 -PI gene. Bases which differ from the α_1 -PI gene are shown. Blank spaces indicate identity, while asterisks indicate insertions. Functional regulatory elements and protein-binding sites discussed in the text are identified below the sequence.

FIG. 2. Transient expression of 5'-end deletion constructs in human and murine cell lines. Constructs for 5'-end deletion analysis are shown at the top. Experimental constructs consisted of segments of the α_1 -PI promoter including 15 bp of exon 1 ligated to the hGRF minigene. The control construct contained ~750 bp of the mouse MTI promoter including 64 bp of exon 1 ligated to the same minigene. A fragment of the α_1 -PI-hGRF construct was subcloned to provide a template for SP6 polymerase from which an antisense RNA probe homologous to the 120- or 105-nucleotide (nt) fragment shown could be generated. Control constructs (3 μ g) and experimental constructs (30 μ g of -522-hGRF or equimolar amounts of deleted constructs) were cotransfected into the hepatoma cell lines HepG2 (human) and BWTG-3 (murine) and into the nonhepatoma cell lines SK-N-SH/SY5Y (human neuroblastoma) and NIH 3T3 (murine fibroblast) by calcium phosphate precipitation. RNase protection was used to quantitate expression of the 5'-end deletion constructs and of the control MTI-hGRF construct as described in Materials and Methods. Experiments which were performed independently are shown separated by blank spaces. The ratio of expression driven by the α_1 -PI promoter constructs to that of the MTI constructs is comparable within a single experiment, while the percentage of maximum expression (-522 construct) for each deletion construct is comparable between experiments (see Table 1).

Figure 2 shows several representative experiments in which 5'-end deletion constructs and the control plasmid were transfected into the human hepatoma cell line HepG2 (26). We also assayed several of the constructs in the murine hepatoma cell line BWTG-3 (48) to determine whether the murine regulatory elements we had shown to be functional in human cells were also active in murine cells. Finally, a subset of the 5'-end deletions was introduced into nonhepatic cell lines—the human neuroblastoma cell line SK-N-SH/SY5Y (2) and the mouse fibroblast line NIH 3T3. It was apparent that sequential deletion of DNA from the 5' end between bases -522 and -80 resulted in a stepwise decrease of expression of the reporter mRNA in hepatoma cells, suggesting the presence of multiple *cis*-acting elements. Significantly, no α_1 -PI promoter-driven expression was detected in nonhepatic cells, although the MTI-driven transcripts were readily apparent. Moreover, the construct containing only 80 bp of upstream sequence continued to display basal tissue-specific expression. By analogy with the human gene, this basal tissue-specific expression is probably driven by the promoter proximal element which binds HNF I. A similar dependence on HNF I for tissue specificity is observed for the α -fetoprotein gene (11).

Table 1 gives the data derived from densitometric scanning of appropriate exposures of from five to eight independent experiments in HepG2 cells. In all cases, data are normalized to the expression of the cotransfected internal standard, MTI-hGRF. Deletion of the entire region from -522 to -133 resulted in a 20-fold decrease in expression from the α_1 -PI promoter, while removal of sequences to -80 reduced expression to barely detectable. Analysis of the expression of the set of 5' deletion mutants in HepG2 cells permitted the localization of several independent *cis*-acting regulatory elements within this region. A 2- to 3-fold decrease in expression occurred between deletion endpoints -318 and -300, and a 4-fold decrease occurred between -197 and -133, while a slight but constant 1.6-fold increase was seen between -242 and -214. The simplest interpretation of these data is that positive *cis*-acting elements are located between -318 and -300 and between -197 and -133, while an inhibitory element may be present between -242/245 and -214. Transfections of the same constructs into murine hepatoma cells (BWTG-3) showed similar results (Fig. 2).

Internal deletion analysis. To understand the functional role of the *cis* elements identified by 5'-end deletions and

TABLE 1. Quantitation of the levels of expression of 5'-end deletion constructs^a

Deletion endpoint	% Expression in the following cell type:		
	Human hepatoma HepG2	Human neuroblastoma SK-N-SH/SY5Y	Mouse fibroblast NIH 3T3
-522	100	ND ^b	ND
-378	93		
-340	74		
-318	72		
-300	31		
-275	26		
-245/-242	15		
-214	24		
-197	20		
-133	5	ND	ND
-80	<1	ND	

^a Appropriate exposures of autoradiograms from RNase protection experiments similar to that shown in Fig. 2 were quantitated by densitometric scanning. All data are normalized for transfection efficiency as described in Materials and Methods. Levels of transcription are expressed as percent expression of the -522 deletion construct. Data for HepG2 cell transfections are derived from at least five independent experiments, while the data for nonhepatic cells were derived from at least two independent experiments. In no case, however, was expression of an α_1 -PI promoter-driven construct observed in a nonhepatic cell type.

^b ND, Not detectable.

TABLE 2. Quantitation of the levels of expression of internal deletion constructs^a

Deletion ends	% Expression in:			
	Hepatoma cells		Nonhepatic cells	
	Human hepatoma HepG2	Mouse hepatoma BWTG-3	Human neuroblastoma SK-N-SH/SY5Y	Mouse fibroblast NIH 3T3
Undeleted (-522)	100	100	ND ^b	ND
-318 to -300	27	40	ND	ND
-318 to -242	30		ND	
-300 to -271	69	64	ND	
-271 to -242	380			ND
-271 to -197	427	610		
-227 to -197	220	440		

^a The data are a compilation of the results of scanning autoradiograms of gels from RNase protection experiments similar to that shown in Fig. 3. Data are normalized as described in Table 1, footnote *a*, and are an average of at least two independent experiments for all transfections.

^b ND, Not detectable.

further define their boundaries, we made a new series of constructs including small internal deletions of putative regulatory regions. All internal deletions were created by using specific PCR-generated upstream fragments containing sequences -522 to the 5' endpoint of the desired deletion and then joining them to the appropriate 5'-end deletion construct described above. This method introduced a 16-

FIG. 3. Internal deletion constructs define functional control elements. A series of internal deletion constructs shown at the top of the figure was derived by ligating cloned and sequenced PCR-generated upstream fragments to 5'-end deletion constructs as described in Materials and Methods. Experimental constructs and the control MTI-hGRF plasmid were cotransfected into four different cell lines as described in the legend to Fig. 2. RNase protection was used to quantitate transient expression of deletion mutants (lower panel). The sequences deleted are shown above each lane. nt, Nucleotides.

nucleotide linker fragment at the joint (see Materials and Methods). Figure 3 and Table 2 show the results of transfecting these deletion mutants into different cell types. It is evident that a positive-acting element, designated HB1 in Fig. 1, lies between -318 and -300 , as was shown by the 5'-end deletion studies. The effect of replacing this element was a reduction to 27% of the undeleted expression level in HepG2 cells or 40% in BWTG-3 cells. In contrast, the deletion of a 26-bp fragment including the AP1 consensus binding sequence located immediately downstream of HB1 at position -287 to -281 resulted in a smaller reduction, to 69% of undeleted expression in HepG2 cells or 64% of wild type in BWTG-3 cells (Fig. 3, lanes 1 and 3). Thus, potential AP1 binding at this site has a less significant effect on expression in the transient assay than HB1 binding. This AP1 site is conserved between mouse and human genes.

The internal deletion studies clearly demonstrated the existence of a negative domain as was suggested by the 5'-end deletions. Three of the constructs in which sequences between -271 and -197 were deleted were expressed at two to six times greater levels than the full-length construct (Fig. 3, Table 2). Moreover, deletion of parts of this domain, either -271 to -242 or -227 to -197 , were each sufficient to cause a twofold or greater increase in expression, so the negative domain may be bipartite. Surprisingly, a construct in which both HB1 and negative element sequences (-318 to -242) were deleted was expressed at approximately the same level as the HB1 deletion alone. This may indicate that the negative element acts to suppress the transcriptional enhancement of the HB1 element. Additional studies are currently under way to further characterize this potential interaction.

Note that the steady-state level of endogenous α_1 -PI mRNA in the mouse hepatoma cell line BWTG-3 is much lower than the level of endogenous α_1 -AT in the human hepatoma line HepG2 (data not shown). This differential expression of endogenous α_1 -PI in these particular cell lines is mirrored in the levels of expression of transfected α_1 -PI promoter-driven constructs. However, the relative effects of deletions of HB1, the negative element, and the positive-acting element between -197 and -133 were the same in both species (Table 2; Fig. 2). Thus, factors capable of regulating the mouse gene are present in hepatic cells of both species. This implies that evolution of *cis*-acting elements is responsible for different patterns of expression of these genes in humans and mice.

Identification of upstream enhancer. A number of genes have been shown to utilize enhancer elements located several kilobases upstream of the transcriptional start site. For example, the liver-abundant gene albumin is modulated by an enhancer located 10 kb upstream of the start site for transcription and α -fetoprotein has been shown to respond to three different upstream enhancers located between 6.5 and 2.5 kb upstream of the start site (20, 40). We therefore initiated a search for *cis*-acting enhancer elements upstream of -522 relative to the α_1 -PI start site. Experimental constructs including 10 kb of DNA upstream of the start site consistently produced two- to threefold enhancement of expression driven by the α_1 -PI promoter (data not shown). To define more accurately the sequences within the 10-kb region responsible for increased transcription, we subdivided the region into three easily resolved restriction fragments which were linked to either the -340 - or -133 -hGRF construct, transfected into hepatoma cells, and assayed for enhancer activity by transient expression assays.

A 1.5-kb fragment of the upstream DNA sequence includ-

FIG. 4. Identification of a far upstream enhancer for murine α_1 -PI-1. The 1.5-kb *Bgl*II fragment extending from -3.6 to -2.1 kb (line A) was inserted in both orientations upstream of the 5'-end deletion -133 -hGRF (line B). The 1.5-kb fragment was further restricted at an internal site (-2.9 kb), and both subfragments were inserted upstream of the 5'-end deletion -133 -hGRF. Only one fragment, which included -3.6 to -2.9 kb, displayed enhancer activity (line C). The fragment was restricted at two internal sites (-3.28 and -3.16 kb, line C), and the three resulting subfragments were inserted into the -133 -hGRF construct. Transient expression and RNase protection as described in the legend to Fig. 2 were used to quantitate the expression of each of the constructs relative to -133 -hGRF. Panel A (below) shows expression of -133 -hGRF compared with expression of equimolar amounts of constructs including the 1.5-kb fragment in either orientation upstream of -133 -hGRF. Panel B shows expression of -133 -hGRF compared with constructs containing the subfragments of the 1.5-kb fragment (line B, above) inserted upstream of -133 -hGRF. Panel C shows expression of -133 -hGRF compared with constructs containing the subfragments of the -3.6 to -2.9 -kb fragment (line C, above) inserted upstream of -133 -hGRF. Mouse MTI-hGRF was cotransfected in all experiments as an internal control for transfection efficiency. nt, Nucleotides.

ing bases -3558 to -2108 was able to enhance expression from the tissue-specific α_1 -PI promoter with 133 bp of proximal 5' sequence (Fig. 4) as well as (to a lesser degree) one which includes 340 bp of proximal sequence (data not shown). This fragment enhanced the test gene in either orientation (Fig. 4, lane 1 compared with lanes 2 and 3) and thus can be operationally defined as an enhancer. Further cleavage and ligation of portions of the 1.5-kb fragment into test constructs resulted in the identification of a 261-bp fragment located at positions -3158 to -2897 which accounted for the majority of the enhancement when linked to the -133 construct (Fig. 4). Additional truncation of the 261-bp fragment destroyed all enhancer activity. A critical protein-binding site may be destroyed by cleavage, or enhancer activity within this fragment may depend on multiple

FIG. 5. Sequence of the 5' end of the upstream enhancer of the transthyretin (TTR) gene (5) is compared with bases -3007 to -3022 of α_1 -PI-I. The AP1 consensus sequence is shown below.

interacting factors. We compared the sequence of this fragment with that of other liver-specific gene upstream enhancers. Interestingly, the 261-bp fragment contains a short region which is very similar to the 5' end of the 100-bp fragment which functions as an upstream enhancer for the transthyretin gene (5) (Fig. 5). While there are partially conserved AP1 (1, 31)-binding sites in both these fragments, it is unlikely that this site is sufficient to account fully for the activity of the 261-bp fragment. The ligation of this AP1 site is a monomer or multimer to the minimal promoter construct (-133-hGRF) has little enhancing activity (data not shown). At least two other sequence elements are involved in the enhancer activity of the 100-bp transthyretin fragment described by Costa et al. (5). Thus, both these upstream enhancers may consist of two or more elements.

The -133-hGRF deletion construct was chosen for testing enhancer activity because it provided the highest level of sensitivity for detection. This was due to the removal of most of the *cis*-acting elements operating on the α_1 -PI-I gene including C/EBP, HB1, and the positive element located between -197 and -133. Note that the -133 construct still maintains a low level of tissue-specific transcription. Enhancers rarely produce an additive effect when linked to an actively transcribed promoter. Not unexpectedly, in analyses in which the upstream enhancer was placed adjacent to either the -522-bp deletion construct or the -340-bp deletion construct, the level of enhancement of transcription was reduced but in no case was it less than twofold (data not shown). If it was true that random sequences added to the -133 construct could somehow result in more efficient binding of a factor which binds near the -133 endpoint, then the insertion of any DNA sequence should increase expression from this construct. However, we observed enhanced expression only when a fragment including bases -3280 to -2897 was inserted.

In summary, analysis of the expression of a series of 5'-end deletions and of more specific internal deletions identified four new functional *cis*-acting DNA sequences. Positive-acting elements lie between -318 and -300 (HB1) and -197 and -133, and a negative-acting element lies between -271 and -197. Deletion of each of these elements had an effect on expression when assayed by transfection into mouse or human cells, which implies that *trans*-acting factors capable of interaction with the mouse elements are conserved between species. In addition, we showed that as few as 80 bp of upstream sequence provide cell-type-specific expression for the murine α_1 -PI-I gene. Finally, we identified a novel upstream element at ~-3 kb which enhances expression of the murine α_1 -PI-I gene in transient expression assays.

Evidence for DNA-binding proteins which recognize *cis*-acting elements. The identification of functional *cis*-acting DNA sequence elements which are involved in the modulation of expression of the α_1 -PI gene implies that specific DNA-binding proteins exist to recognize the elements. We

used various combinations of DNase I footprinting (14), methylation interference (46), organic copper footprinting (30), and gel retardation (12, 15) analysis to identify the existence of such binding proteins for three of the elements described above. The specific *cis*-acting elements at positions -3158 to -2897 (upstream enhancer), -318 to -300 (HB1), and -272 to -197 (negative element) were analyzed as described above. Small DNA fragments which included each region were isolated, labeled, and subjected to footprinting analysis independently.

The 132-bp fragment spanning the region from -420 to -294 was end labeled on both strands and used in methylation interference studies of the HB1 element (-318 to -300) (46). Several G residues on the noncoding strand (at -318, -317, -314, -313) were protected (Fig. 6A). When the same probe was used for organic copper footprinting, a region of protection on both strands extending from -313 to -300 could be seen (Fig. 6B). These experiments confirmed the presence of proteins in liver cell nuclei which can specifically recognize and bind to DNA sequences in the region from -318 to -300, which corresponds to the functional positive-acting element HB1.

Analysis of the negative element (-272 to -197) was more problematic owing to multiple, closely spaced protein-binding sites. For example, partially purified C/EBP binds to positions -209 to -185, HNF 3 binds to positions -199 to -181 (3), and a consensus AP1-binding site, TGACTCA, appears between -287 and -281. Nevertheless, a fragment including the region from -294 to -197 was subjected to methylation interference analysis and showed protection of G residues at positions -254 and -255 on the sense strand (Fig. 6A). This implies that there is a specific interaction of protein with this sequence. The results of copper footprinting and DNase I footprinting with crude nuclear extracts in this region were more equivocal because the region interacts with a number of different binding proteins. By using decreasing amounts of crude liver nuclear protein in a DNase I footprinting assay, it was possible to resolve a number of potential specific interactions in the region from -340 to -170 (Fig. 6C). At the highest liver protein concentrations, the entire region from -300 to -180 was protected from digestion, which may be anticipated based on the number of proteins which appear to bind to this region. As the concentration of crude nuclear extract decreased, several discrete protected regions appeared on each strand in a symmetrical pattern. The furthest 5' protected sequence was located between -287 and -281 and corresponded to the consensus AP1-binding site (Fig. 6C, lanes 2 and 3). We showed that specific deletion of this sequence has an effect on the level of expression from the α_1 -PI promoter (Fig. 3; Table 2) but does not reduce expression as much as deletion of the adjacent HB1 element. This points out the importance of demonstrating function before evaluating the significance of observed DNA-binding activities. The next distinct DNase I-protected sequence extended from -257 to -240 (Fig. 6C, lanes 2 and 3). This region included the protected G residues seen in the methylation interference studies (Fig. 6A) and corresponded to the 5' half of the negative element. Thus, methylation interference and DNase I protection provide evidence for protein-DNA interaction in the region which is functionally defined as a negative element. An additional well-demarcated DNase I-protected sequence could be seen at positions -215 to -185, where C/EBP (-210 to -185) and HNF 3 (-199 to -181) have been shown to bind (3). 5'-end deletion in this region (-197 to -133) resulted in decreased expression from the α_1 -PI promoter, consistent with a role of

FIG. 6. DNA-protein binding assays show that liver-derived nuclear proteins interact with functional *cis* regulatory elements. Three different assays were performed on several fragments as described in the text and in Materials and Methods. (A) Methylation protection of G residues in two fragments. On the left, the fragment extends from -294 to -197 and includes the functionally defined negative element (NE). Two G residues were protected from methylation at -254 and -255. On the right, the fragment extends from -420 to -294 and includes the positive element HB1. Four G residues are protected from methylation at positions -318, -317, -314, and -313. Bound refers to probe which is gel retarded owing to protein interaction. Unbound refers to DNA probe alone (see Materials and Methods). (B) Copper footprint of part of the HB1 fragment described above. The protected region extends from -313 to -300 on both strands. (C) DNase I footprint on a fragment extending from -340 to -170. From 14 to 70 μ g of protein extract was used per reaction. M lanes contain a G ladder, while C lanes contain labeled DNA incubated without extract. At the lowest extract concentration (lanes 3 and 10), distinct footprints can be seen on both strands, corresponding to the binding sites for the *trans*-acting factors identified on the left: C/EBP-HNF 3, AP1, and the region (NE) identified as a functional negative element in this report. (D) Initial footprint analysis of part of the functionally defined upstream enhancer, including bases -3007 to -2897. Samples of 75 and 100 μ g of protein were used in the lanes shown. The protected region extends from -2998 to -2980.

either C/EBP or HNF 3 in positive regulation of α_1 -PI expression.

We also performed initial DNase I footprint analysis on the 261-bp upstream enhancer fragment (-3158 to -2897) and identified a region which is well protected from DNase I digestion at positions -2998 to -2981 (Fig. 6D). The protected region includes the sequence (GTGATCA) which is partially homologous to the AP1 consensus sequence (GTGACTCA). This region is necessary but not sufficient for enhancer activity since constructs in which from one to seven of these AP1 sites are linked to the test gene failed to show any enhancement of transcription. On the other hand,

deletion of the AP1 site inactivates the enhancer activity of the 261-bp enhancer fragment (data not shown). It appears that additional sites within the fragment are necessary for the enhancer function. Note that the liver-abundant gene, transthyretin, contains a similar imperfect AP1 consensus element in its far upstream enhancer element (Fig. 5) which, when deleted, resulted in partial inactivation of the 100-bp long element.

Tissue specificity of DNA-binding proteins. Basal tissue-specific expression of the α_1 -PI-I gene is controlled by promoter-proximal elements. However, it is important to determine whether any of the elements which we identified

upstream of the HNF II protein-binding site (–125 to –101), which increase the rate of transcription in liver, are also recognized by tissue-specific DNA-binding proteins. This was done by performing gel retardation analyses with nuclear protein extracts from several different tissue sources which differ in their expression of endogenous α_1 -PI mRNA. Since there is a nearly perfect AP1 consensus binding site in the far upstream enhancer element, the gel retardation analysis of protein binding to this region was found to be inconclusive since it was impossible to determine, by a simple binding assay, which AP1-like protein complex actually binds to this sequence from a particular tissue. We therefore focused our analysis on the HB1 element and the negative element.

Protein extracts were prepared from liver, kidney, and 3T3 cell nuclei and tested for the presence of sequence-specific DNA-binding proteins in a gel retardation assay. The labeled DNA probe used for the HB1 element spanned the sequence –325 to –287 and did not include the consensus AP1 element immediately downstream of HB1 (Fig. 1). The probe for the negative element spanned the region from –269 to –227 and did not include any sequences previously shown to bind to nuclear proteins. Both fragments bound to specific proteins from liver and kidney (Fig. 7) as shown by the ability of an excess of specific cold competitor to abolish DNA binding of a subset of the shifted bands for each fragment. Interestingly, no HB1-specific DNA-binding activity could be seen in the 3T3 cell extracts (Fig. 7, upper panel). These data are consistent with the observed pattern of expression of α_1 -PI in that both liver and kidney express endogenous α_1 -PI mRNA while 3T3 cells do not. Negative element-binding activity (Fig. 7, lower panel), however, was seen in all three tissues, showing that this activity is not tissue specific and that the 3T3 extracts used contain binding activity. Since the negative element-binding activity was present in tissues which express α_1 -PI, we must conclude that either negative element binding alone is not sufficient to completely inhibit expression of the gene in the presence of positive-acting factors or tissue-specific positive factors can prevent binding of the negative factor in vivo.

DISCUSSION

In this report, we identified and characterized four new functional *cis*-acting DNA elements located upstream of the liver-abundant α_1 -PI-I gene. Two of the elements were positive effectors of transcription when tested by transient transfection assay into human and mouse hepatoma cells. A third element inhibited transcription of the murine α_1 -PI-I promoter in liver cells. The fourth element, which was located approximately 3 kb upstream of the start site of transcription and had a short region of homology to the upstream enhancer of the liver-abundant gene encoding transthyretin (3), behaved as an enhancer in that it worked in a position- and orientation-independent manner. However, the effects of this enhancer were not entirely additive when it was placed upstream of additional positive-acting elements.

Each of the elements appeared to interact with one or more DNA-binding proteins in liver nuclear extracts and thus is likely to act via *trans*-acting transcriptional factors. Although others have identified sites within the murine α_1 -PI-I promoter region which interact with known DNA-binding proteins, our studies concentrated on elements for which function can be demonstrated in a transfection assay. Perhaps because of our use of a large number of 5'-end deletions as well as internally deleted constructs in these assays, we identified novel regulatory elements. Initial stud-

FIG. 7. Gel retardation analysis of the HB1 and negative *cis*-acting elements. End-labeled DNA fragments corresponding to the HB1 element (–325 to –287) or negative element (–269 to –227) were incubated with 20 μ g of nuclear extract per ml from liver, kidney, or 3T3 cells in the presence of poly(dI-dC) (100 μ g/ml) and no additional carrier DNA, a 30-fold excess of specific competitor, or a 30-fold excess of nonspecific competitor as described in Materials and Methods. The specific competitors for the HB1 element and negative elements were synthesized by the PCR. Specific competitor for the HB1 element extended from –340 to –289, and competitor for the negative element extended from –271 to –228. The nonspecific competitor HB1 was the negative element fragment and that for the negative element was the HB1 fragment. The presence (+) or absence (–) of extract (Ext.), specific competitor (Sp.), or nonspecific competitor (N.Sp.) is indicated above each lane. The small arrows indicate the positions of gel-retarded fragments which can be inhibited by the presence of specific competitor.

ies by gel retardation analyses on the proteins which specifically bind to the negative element and to HB1 showed that HB1-binding activity can be detected only in tissues which express α_1 -PI-I, whereas negative element-binding activity appears to be ubiquitous.

We compared two murine gene promoter sequences with that of the human gene, with special emphasis on the sequence domains which have been shown by others to have functional significance for the regulation of the human gene (8, 42) or which we find to be significant in the regulation of expression of the mouse gene. This comparison showed that the DNA sequences have diverged in several key regions,

perhaps accounting for the differences in expression patterns seen in the two species. On the other hand, we found that the murine gene is regulated in a human cell line in the same way as in a mouse cell line, suggesting that while the *cis*-acting DNA elements have evolved during species divergence, at least some of the relevant regulatory protein factors have retained their activity. This is further supported by the observation that the human gene follows the human pattern of expression when expressed in transgenic mice (25, 43). The existence of a five-member gene family in the mouse versus a single active human gene may result in α_1 -PI serum proteins with significantly different characteristics, such as their substrate specificity or turnover rate in serum. This may explain how the mouse and human can use different patterns of expression to effect similar antiprotease function.

This report presents the first clear evidence for a negative element active in liver cells within the murine α_1 -PI-I gene. Interestingly, a comparison of the levels of expression from the 5'-end deletion constructs -275 and -242 showed a minor (<10%) decrease in promoter activity, while internal deletion of the same sequences, leaving the upstream region intact, resulted in a 3.8-fold activation of the promoter (Fig. 3; Table 2). Our interpretation of these data is that factors binding to the negative element must interact with positive regulatory elements, such as HB1, further upstream and thereby inhibit upstream enhancing activity. The fact that a construct in which bases -318 to -242 are deleted (i.e., HB1 and part of the negative element) is expressed at approximately the same level as a deletion of sequences from -318 to -300 (i.e., HB1 only) supports the idea that the proteins which bind to the negative element and HB1 interact in some way. The removal of negative element sequences in the absence of HB1 did not result in the loss of repression.

The independent deletion of either of two nonoverlapping internal segments (-271 to -242 or -227 to -197) gave similar levels of derepression, which were somewhat lower than the level observed when the entire region (-271 to -197) was deleted. Thus, the entire domain is important for negative regulation of the gene, either due to the binding of a single protein most effectively to the entire region or due to two or more binding sites for negative factors in which their effects are at least partially additive. The region (-272 to -197) is G+T-rich (>68%), and between -255 and -215, the sequence is nearly 90% G+T. Whether or not this unusual DNA sequence is significant to the repressing activity is not known. Note that we present footprinting evidence that specific protein binding occurs throughout this region (Fig. 6), and its biological activity is likely to be the result of multiple protein interactions. It is unlikely that the effects of the internal deletion of the negative element are due to artifacts caused by altering the spacing of positive elements, because different-sized deletions within the same region produced very similar results (Fig. 3; Table 2).

The tissue-specific expression of the α_1 -PI gene is clearly the result of complex interactions of specific and general transcription factors acting in concert. Our data indicate that promoter-proximal elements are necessary and sufficient for basal expression of the gene in hepatoma cells. Even transfection of a construct from which all sequences through the CAAT box to position -80 are deleted resulted in barely detectable but tissue-specific expression. This is most likely due to the binding of HNF I within this region (11, 37). Significantly, we never observed transcription in a nonhepatic cell type from any construct which contained this promoter-proximal segment of DNA despite the presence of a consensus TATA element. Thus, HNF I binding may be required for

all promoter activity in this gene. The presence or absence of promoter elements upstream of -80 appears only to affect the quantitative level of expression in these constructs and not the ability of the α_1 -PI promoter to be expressed in a nonhepatic cell. Therefore, it would be incorrect to attribute observed tissue-specific expression to any of the upstream elements. Nevertheless, it is important to determine whether the proteins which bind to the upstream elements are present only in liver cells or in other tissues as well.

Gel retardation experiments showed that the negative element is recognized by DNA-binding proteins which are relatively ubiquitous, while the HB1 element is recognized by proteins only found in tissues which express α_1 -PI (i.e., liver and kidney but not fibroblasts). Although it is difficult to draw conclusions about the reason for these patterns of expression, it is attractive to hypothesize that the negative element proteins ensure that transcription of the α_1 -PI genes is turned off in inappropriate tissues. It is, however, significant that deletion of the negative element does not result in expression of the α_1 -PI promoter in fibroblasts or neuroblastoma cells (Fig. 2 and 3). This can be interpreted in a variety of ways. One is that the negative element only suppresses transcription in the presence of upstream positive elements and that most of these are tissue specific. Alternatively, one can hypothesize that in a transient expression assay, low-level expression from the promoter cannot be detected, but that in vivo the negative element is required to achieve a truly silent promoter. To address this issue directly, we are assessing the function of internal deletions of the negative element in transgenic mice in which it might be possible to observe low-level expression in inappropriate tissues.

DeSimone and co-workers (8, 9) have recently identified two different negative regulatory activities in the human α_1 -AT gene. They showed that spleen contains a factor, LF-B2, which footprints the HNF I (LF-B1) region and which might be an inhibitor since its binding prevented binding of LF-B1. More recently, they have described a human DNA sequence required for negative regulatory activity located where HNF II (or LF-A1) binds. The activity is found in HeLa cells and is thought to be at least partially responsible for the liver-specific expression of α_1 -AT. Note that the characteristics of the negative activity in the human gene are quite different from those we see in the mouse, with respect to both location relative to the start site for transcription and the sequences themselves. Furthermore, the human negative activity is found only in nonhepatic cells, while we found negative activity in hepatic cells as well. Thus, although we cannot exclude the possibility that the negative regulatory mechanisms seen in the human and mouse systems are related, we think it is unlikely.

While our data demonstrate that the extremely G+T-rich domain located between -272 and -197 in the mouse can repress transcription in both human and murine hepatic cells, this negative-acting domain corresponds to a region in the human gene (-261 to -210) which has been characterized as a strong positive-acting regulatory region. Deletion of this region from the human gene results in a four- to five-fold reduction of expression of human α_1 -AT. It has been proposed that the functional element in the human gene may be a core enhancer sequence, TGTGTTT (8, 42). Comparison of the murine and human sequences (Fig. 1) reveals that both of the mouse genes contain the sequence TTTGTTGT, which effectively interrupts the core enhancer consensus (47) and results in the formation of the perfect direct repeat, -251 GTTGTGttt gttgtGTTT -234. Although the significance of this repeat is unknown, our data showed that the enhancer

function of the core sequence is completely inactivated by this insertion (Table 1, deletions -275 through -214).

On the other hand, our data showed that an important positive regulatory activity for the mouse gene is located between -318 and -300, designated HB1. Sequence comparison of the mouse and human genes as aligned in Fig. 1 shows only ~50% homology for this region. In contrast to our data, DeSimone et al. (8), studying the human gene, found only a small drop in transcriptional efficiency between 5'-end deletion constructs including 337 or 261 bp of upstream sequence (less than a 5% drop [Fig. 2A in reference 8]), using transient chloramphenicol acetyltransferase assays in HepG2 cells. So the HB1 element domain of the human gene is not functional as an enhancer of transcription in this very similar assay system.

The identification of a novel far upstream enhancer element for the murine α_1 -PI-I gene is interesting in light of similar findings for other liver-abundant genes (4, 20, 40). Comparison of the far upstream enhancer DNA sequence of the α_1 -PI-I gene (-3.16 kb) with the far upstream enhancer of the transthyretin gene (-1.9 kb) revealed a small region of striking sequence similarity (Fig. 5). Within both enhancers, there is a highly conserved AP1-binding motif, but the AP1 motif alone is not sufficient for enhancement of the α_1 -PI promoter. It is attractive to speculate that the enhancer requires the interaction of several different DNA-binding proteins. Although the DNA footprint of this region with crude nuclear extract showed only a single binding site within the 261-bp active fragment, further analysis with partially purified protein may reveal additional sites. It is significant that the upstream enhancer does not appear to function in transient assay in BWTG-3 mouse hepatoma cells. Since mouse liver had proteins which can bind to the functional region and since BWTG-3 cells poorly express α_1 -PI constructs generally, we hypothesize that these cells simply do not have the necessary set of proteins required for enhancer function. A direct test of the upstream enhancer in transgenic mice is under way and will resolve this issue much more definitively.

The conclusions drawn from our experiments regarding the functional *cis*-acting elements regulating α_1 -PI expression differ from those reported by Grayson et al. (18) in several respects. By searching a greater distance upstream of the promoter, our analysis uncovered a far upstream enhancer which their analysis did not, while they showed that the region from -522 to -168 can also act as an orientation-independent enhancer. The level of enhancement of these two elements, separated in the genome by ~2.5 kb, is equivalent (~7- to 10-fold). However, they do not seem to act additively, since our data showed that the far upstream enhancer is less potent when linked to constructs including part or all of the region from -522 to -168. Transgenic studies under way will show whether these observations are also true in vivo. In addition, our use of a combination of 5'-end and internal deletion constructs revealed the negative element, which their studies did not, and showed that a significant positive-acting element not identified in their work, HB1, is found between -318 and -300.

The deletion of the HB1 element resulted in an approximately threefold reduction in expression (Table 1). In the experiments of Grayson et al. (18), the primary element which enhances transcription of this gene lies between -457 and -397 with respect to the start site for transcription. This conclusion was derived solely from a comparison of the levels of expression of 5'-end deletion constructs which included 457, 397, 345, or 294 bp upstream with an approx-

imately fourfold reduction of transcription seen between them. In contrast, we found that deletion from -522 to -407 or -378 results in only a small reduction in transcriptional efficiency, to no less than 93%. Moreover, deletion to -318 still yielded 72% of the expression seen with the -522 construct. To determine whether subtle differences in our constructs and reporter genes could account for the discrepancy, we prepared two minigene constructs which were very similar to the constructs used by Grayson et al. (18) and compared expression of -522 and -340-base deletion constructs in HepG2 cells. Our results showed that this deletion on the minigene reporter sequence still only yielded an ~72% level of expression (data not shown). We repeated our 5'-end deletion studies on the controversial segment more than five times in HepG2 cells and at least twice in mouse BWTG-3 cells with very consistent results. Therefore, we must conclude that the differences must be cell specific and must relate to subtle differences in cells grown in our two laboratories.

Note that the choice of cell line for transfection analysis is obviously extremely important. Before deciding on using HepG2 and BWTG-3 hepatoma cells, we tested four other rodent hepatoma cell lines for their ability to express α_1 -PI gene constructs with entirely negative results despite the fact that in each case, the tested cells were reported to express several different liver-specific gene products. This implies that cell lines can evidently change with respect to the expression of important *trans*-acting factors which are required for the expression of individual genes. The best available assay for the importance of *cis*-acting elements in the tissue-specific expression of a particular gene is to test the effects of deletions in transgenic mice. Therefore, we initiated this analysis for a number of the elements which regulate expression of α_1 -PI-I in cultured cells. In this way, perhaps, we can resolve important questions about the role of the upstream enhancer, the HB1 element, and the negative element in normal mouse liver gene expression.

ACKNOWLEDGMENTS

We thank June Biedler of the Sloan-Kettering Institute for Cancer Research for the human neuroblastoma cell line SK-N-SH/SY5Y.

This work was supported by Public Health Service grants CA 39553 and CA 31855 (to K.S.K.) and CA 30117 (to L.R.) from the National Institutes of Health. J.T. was supported by MSTP grant T32GM7288-13, and K.T.M. was supported by a National Cancer Institute National Research Service award (5 F32 CA 08385).

LITERATURE CITED

1. Angel, P., M. Imagawa, R. Chiu, B. Stein, R. Imbra, R. J. Imbra, H. J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin. 1987. Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. *Cell* 49:729-739.
2. Biedler, J. L., L. Helson, and B. A. Spengler. 1978. Multiple neurotransmitter synthesis by human neuroblastoma cell lines and clones. *Cancer Res.* 38:3751-3757.
3. Costa, R. H., D. R. Grayson, and J. E. Darnell, Jr. 1989. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α_1 -antitrypsin genes. *Mol. Cell. Biol.* 9:1415-1425.
4. Costa, R. H., E. Lai, and J. E. Darnell, Jr. 1986. Transcriptional control of the mouse prealbumin (transthyretin) gene: both promoter sequences and a distinct enhancer are cell specific. *Mol. Cell. Biol.* 6:4697-4708.
5. Costa, R. H., E. Lai, D. R. Grayson, and J. E. Darnell, Jr. 1988. The cell-specific enhancer of the mouse transthyretin (prealbumin) gene binds a common factor at one site and a liver-specific factor(s) at two other sites. *Mol. Cell. Biol.* 8:81-90.
6. Courtois, G., J. G. Morgan, L. A. Campbell, G. Fourel, and

- G. R. Crabtree. 1987. Interaction of a liver-specific nuclear factor with the fibrinogen and α_1 -antitrypsin promoters. *Science* 238:688-692.
7. Derman, E., K. S. Krauter, L. Walling, C. Weinberger, M. Ray, and J. E. Darnell, Jr. 1981. Transcriptional control of liver specific mRNAs. *Cell* 23:731-739.
8. DeSimone, V., G. Ciliberto, E. Hardon, G. Paonessa, F. Palla, L. Lundberg, and R. Cortese. 1987. *cis* and *trans*-acting elements responsible for the cell-specific expression of the human α_1 -antitrypsin gene. *EMBO J.* 6:2759-2766.
9. DeSimone, V., and R. Cortese. 1989. A Negative regulatory element in the promoter of the human α_1 -antitrypsin gene. *Nucleic Acids Res.* 17:9407-9415.
10. Erlich, H. A., D. H. Gelfand, and R. S. Saiki. 1988. Specific DNA amplification. *Nature* (London) 331:461-462.
11. Feuerman, M. H., R. Godbout, R. S. Ingram, and S. M. Tilghman. 1989. Tissue-specific transcription of the mouse α -fetoprotein gene promoter is dependent on HNF1. *Mol. Cell. Biol.* 9:4202-4212.
12. Fried, M., and D. M. Crothers. 1981. Equilibria and kinetics of Lac repressor operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* 9:6505-6525.
13. Gadek, J. E., and R. G. Crystal. 1982. α_1 -Antitrypsin deficiency, p. 1450-1467. In J. B. Stanbury, J. B. Wyngaarden, D. S. Frederickson, J. L. Goldstein, and M. S. Brown (ed.), *The metabolic basis of inherited disease*. McGraw-Hill Book Co., New York.
14. Galas, D., and A. Schmitz. 1978. DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Res.* 5:3157-3170.
15. Garner, M. M., and A. Revzin. 1981. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: applications to components of the *E. coli* lactose operon regulatory system. *Nucleic Acids Res.* 9:3074-3090.
16. Graham, F., and A. van der Eb. 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52:456-457.
17. Grayson, D. R., R. H. Costa, G. Xanthopoulos, and J. E. Darnell, Jr. 1988. One factor recognizes the liver-specific enhancers in α_1 -antitrypsin and transthyretin genes. *Science* 239:786-788.
18. Grayson, D. R., R. H. Costa, K. G. Xanthopoulos, and J. E. Darnell, Jr. 1988. A cell-specific enhancer of the mouse α_1 -antitrypsin gene has multiple functional regions and corresponding protein binding sites. *Mol. Cell. Biol.* 8:1055-1066.
19. Hammer, R. E., R. L. Brinster, M. G. Rosenfeld, R. M. Evans, and K. E. Mayo. 1985. Expression of human growth hormone-releasing factor in transgenic mice results in increased somatic growth. *Nature* (London) 315:413-416.
20. Hammer, R. E., R. Krumlauf, S. A. Camper, R. L. Brinster, and S. M. Tilghman. 1987. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. *Science* 235:53-58.
21. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene* 28:351-359.
22. Hill, R. E., P. H. Shaw, R. K. Barth, and N. D. Hastie. 1985. A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts. *Mol. Cell. Biol.* 5:2114-2122.
23. Johnson, D., and J. Travis. 1979. The oxidative inactivation of human α_1 -proteinase inhibitor: further evidence for methionine at the reactive center. *J. Biol. Chem.* 254:4022-4026.
24. Johnson, P. F., W. H. Landschulz, B. J. Graves, and S. McKnight. 1987. Identification of a rat liver protein that binds to the enhancer core element of three animal viruses. *Genes Dev.* 1:133-146.
25. Kelsey, G. D., S. Povey, A. E. Bygrave, and R. H. Lovell-Badge. 1987. Species- and tissue-specific expression of human α_1 -AT in transgenic mice. *Genes Dev.* 1:161-171.
26. Knowles, B. B., C. C. Howe, and D. P. Aden. 1980. Human hepatocellular carcinoma cell lines secrete the major plasma proteins and hepatitis B surface antigen. *Science* 209:497-499.
27. Koopman, P., S. Povey, and R. Lovell-Badge. 1989. Widespread expression of human α_1 -antitrypsin in transgenic mice revealed by *in situ* hybridization. *Genes Dev.* 3:16-25.
28. Kraft, R., J. Tardiff, K. S. Krauter, and L. A. Leinwand. 1988. Using miniprep plasmid DNA for sequencing double stranded templates with sequenase. *BioTechniques* 6:544-547.
29. Krauter, K. S., B. A. Citron, M.-T. Hsu, D. Powell, and J. E. Darnell, Jr. 1986. Isolation and characterization of the alpha-1 antitrypsin genes of mice. *DNA* 5:29-36.
30. Kuwabara, M. D., and D. Sigman. 1987. Footprinting DNA-protein complexes *in situ* following gel retardation assays using 1,10-phenanthroline-copper ion: *E. coli* RNA polymerase-lac promoter complexes. *Biochemistry* 26:7234-7238.
31. Lee, W., P. Mitchell, and R. Tjian. 1987. Purified transcription factor Ap1 interacts with TPA-inducible enhancer elements. *Cell* 49:741-752.
32. Li, Y., R.-F. Shen, S. Y. Tsai, and S. L. C. Woo. 1988. Multiple hepatic *trans*-acting factors are required for *in vitro* transcription of the human α_1 -antitrypsin gene. *Mol. Cell. Biol.* 8:4362-4369.
33. Lichtsteiner, S., J. Wuarin, and U. Schibler. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. *Cell* 51:963-973.
34. Long, G. L., T. Chandra, K. Kurachi, S. L. C. Woo, and E. W. Davie. 1984. Complete sequence of the cDNA for human α_1 -AT and the gene for the S variant. *Biochemistry* 23:4828-4837.
35. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
36. Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035-7056.
37. Monaci, P., A. Nicosia, and R. Cortese. 1988. Two different liver-specific factors stimulate *in vitro* transcription from the human α_1 -antitrypsin promoter. *EMBO J.* 7:2075-2087.
38. Morrow, B., S. Johnson, and J. Warner. 1989. Yeast rDNA enhancer binding proteins. *J. Biol. Chem.* 264:9061-9068.
39. Parker, C. S., and J. Topol. 1984. A Drosophila RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. *Cell* 36:357-369.
40. Pinkert, C. A., D. M. Ornitz, R. L. Brinster, and R. D. Palmiter. 1987. An albumin enhancer located 10 kb upstream functions along with its promoter to direct efficient, liver-specific expression in transgenic mice. *Genes Dev.* 1:268-276.
41. Powell, D. J., J. M. Friedman, A. J. Oulette, K. S. Krauter, and J. E. Darnell, Jr. 1984. Transcriptional and post-transcriptional control of specific messenger RNAs in adult and embryonic liver. *J. Mol. Biol.* 179:21-35.
42. Shen, R., Y. Li, R. N. Sifers, H. Wang, C. Hardick, S. Y. Tsai, and S. L. C. Woo. 1987. Tissue-specific expression of the human alpha-1 antitrypsin gene is controlled by multiple *cis*-acting regulatory elements. *Nucleic Acids Res.* 15:8399-8415.
43. Sifers, R., J. Carlson, S. Clift, F. DeMayo, D. W. Bullock, and S. L. C. Woo. 1987. Tissue specific expression of the human alpha-1-antitrypsin gene in transgenic mice. *Nucleic Acids Res.* 15:1459-1475.
44. Struhl, K. (ed.). 1989. *Current protocols in molecular biology*, p. 4.1.2. Greene Publishing Associates and John Wiley & Sons, Inc., New York.
45. Travis, J., and D. Johnson. 1978. Structural evidence for methionine at the reactive site of human α_1 -protease inhibitor. *J. Biol. Chem.* 253:7142-7144.
46. Treisman, R. 1986. Identification of a protein-binding site that mediates transcriptional response of the *c-fos* gene to serum. *Cell* 46:567-574.
47. Weiher, H., M. König, and P. Gruss. 1983. Multiple point mutations affecting the SV40 enhancer. *Science* 219:626-631.
48. Widen, S. G., and J. Papaconstantinou. 1986. Liver-specific expression of the mouse α -fetoprotein gene is mediated by *cis*-acting DNA elements. *Proc. Natl. Acad. Sci. USA* 83:8196-8200.